Research group: Nataša Jonoska, Masahico Saito, Johnathan Burns, Daniel Cruz, Daria Karpenko, Denys Kukushkin, Rick Wallace, Micah Wine

Rearrangement Pathways in Scrambled DNA Sequences

Richard Wallace and Maja Milošević

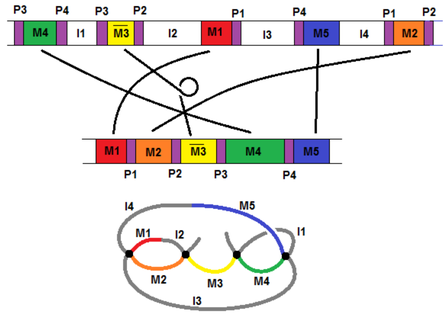
USF Department of Mathematics and Statistics

Abstract

We developed a program that determines possible pathways in which the homologous DNA rearrangement occurs by using a modified version of the depth first search algorithm. This program takes genetic sequences as inputs and outputs the possible rearrangement pathways. This program will be applied to a massively scrambled ciliate genome.

Mathematical Description

During DNA rearrangement, scrambled DNA needs to be rearranged and have certain parts cut out, as shown below:



The model represents the data using graphs. An example is shown to the right. In this graph, the edges represent the DNA sequences and the vertices represent the spot in which they overlap before rearrangement. In reality, not all of the sequences are flexible enough to form cyclic molecules, and this program calculates which ones can.

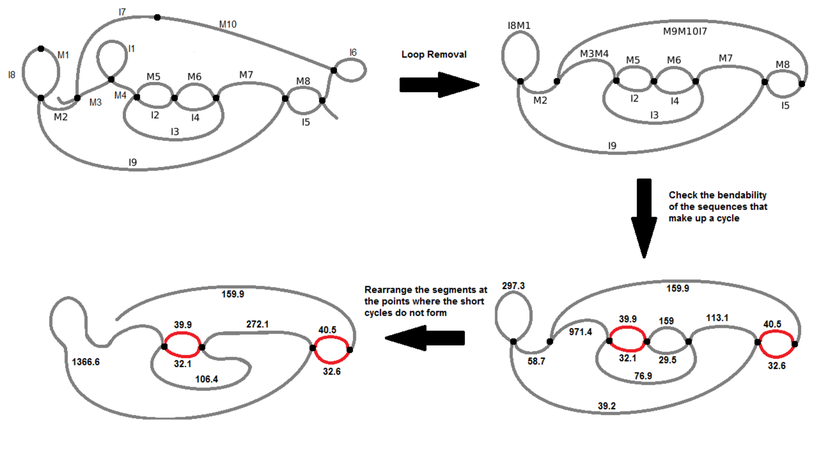
Rearrangement

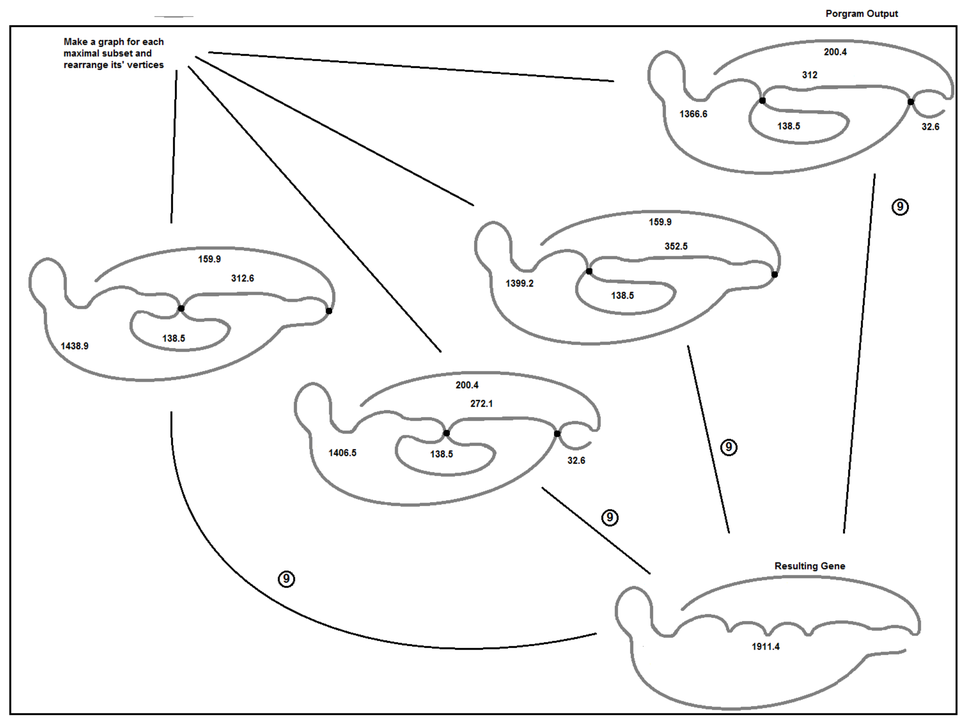
During rearrangement, the DNA is expected to fold back onto itself. This makes the bendability of the DNA a very imporant factor. The program takes into account the following factors to calculate how bendable the DNA are (according to [3]):

- The length of the sequence to be bent.
- The nucleotide content of the sequence (the bases nearest neightbor bendability factor).

Output Step-by-Step

This is the output of Actin 1 gene of *Oxytricha Trifallax*





Algorithm

The following is a breakdown of the algorithm the program uses:

1. INPUT: This program takes an annotated precursor DNA sequence.
2. Convert sequences into graph.
3. Remove all loops of excised data and the first and last vertex (according to data, this is differnet than the rearrangement process) [1]
4. Find all cycles in the graph.
5. Apply length/sequence paramter to cycles [3]
6. For each small cycle find maximal subsets of its vertices where recombination can occur [3]
7. For each vertex not in any subset, apply recombination at the vertex
8. Make a new graph for each maximal subset and apply recombination at those vertices
9. Repeat (2) - (8) for each newly formed graph
10. Output the graphs in a tree structure

References

[1] M. Mollenbeck, Y. Zhou, ARO. Cavalcanti, F. Jonsson, BP. Higgins, et al. (2008) The Pathway to Detangle a Scrambled Gene. PLoS ONE 3(6): e2330. doi:10.1371/journal.pone.0002330.
[2] R.L. Baldwin, J. Langowski, D. Shore, Formation of small circular DNA molecules via an in vitro site-specifc recombination system, Proc. Natl. Acad. Sci. (USA) 78 (1981) 4833-4837.13
[3] K. Abremski, R. Hoess, A. Wierzbicki, Formation of small circular DNA molecules via an in vitro site-specific recombination system, Gene 40 (1985) 325-329.

Acknowledgements

Supported by NSF grants CCF-1117254 and DMS-0900671
and NIH grant R01GM109459-01

Contact Information

- Web: http://knot.math.usf.edu
- Email: rbwallace@mail.usf.edu